

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2003年 3月31日

出 願 番 号
Application Number:

特願2003-093118

[ST.10/C]:

[JP2003-093118]

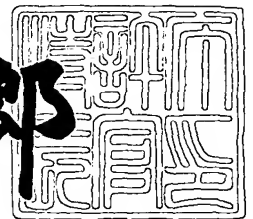
出 願 人
Applicant(s):

株式会社日立製作所

2003年 6月 9日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3044543

【書類名】 特許願

【整理番号】 K02009481A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 13/10

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

 【氏名】 武田 貴彦

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

 【氏名】 伊東 隆介

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

 【氏名】 海谷 佳一

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

【手数料の表示】

 【予納台帳番号】 013088

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

【物件名】 要約書 1
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶システム及びその制御方法

【特許請求の範囲】

【請求項 1】

上位装置から受領した記憶装置に対するデータの書き込みや読み出しを制御する
第 1 の記憶制御装置を有する記憶システムであって、

前記第 1 の記憶制御装置は、前記上位装置に接続され、前記上位装置からのコマンドを処理する第 1 の処理部と、

前記上位装置から受領したデータを一時的に保持するキャッシュメモリと、

前記記憶システムの管理情報を保持するメモリと、

前記キャッシュメモリ上のデータの前記記憶装置に対する書き込みを制御する
第 2 の処理部と、を有し

前記記憶システムは、前記第 2 の処理部に、第 2 の記憶制御装置が接続されることを特徴とする記憶システム。

【請求項 2】

請求項 1 に記載の記憶システムにおいて、前記第 1 の記憶制御装置が、前記第 1 の記憶制御装置の制御下にある第 1 の記憶装置上の論理デバイスと、前記第 2 の記憶制御装置の制御下にある第 2 の記憶装置上の論理デバイスとの間にペア関係を生成することを特徴とする記憶システム。

【請求項 3】

請求項 2 に記載の記憶システムにおいて、前記メモリの管理情報として、第 1 及び第 2 の記憶制御装置の制御情報、デバイス情報、差分管理情報を含むことを特徴とする記憶システム。

【請求項 4】

請求項 1 に記載の記憶システムにおいて、前記第 2 の記憶制御装置は、この記憶制御装置に制御される記憶装置上に少なくとも 2 以上の論理デバイスを有し、前記第 1 の記憶制御装置の前記メモリ内に、前記 2 以上の論理デバイス同士のペア状態を管理する管理情報を保持することを特徴とする記憶システム。

【請求項 5】

請求項 4 に記載の記憶システムにおいて、

前記管理情報は、前記第 2 の記憶制御装置によって制御される記憶装置上の記憶デバイスのペア状態が、スプリット状態である場合に、その差分情報を保持する差分管理情報を有することを特徴とする記憶システム。

【請求項 6】

上位装置から受領したデータの記憶装置に対する書き込みや読み出しを制御する第 1 の記憶制御装置を有する記憶システムであって、

前記第 1 の記憶制御装置は、前記上位装置に接続され、前記上位装置からのコマンドを処理する第 1 の処理部と、

前記上位装置から受領したデータを一時的に保持するキャッシュメモリと、

前記記憶システムの管理情報を保持するメモリと、

前記キャッシュメモリ上のデータの前記記憶装置に対する書き込みを制御する第 2 の処理部と、を有し

前記記憶システムは、前記第 2 の処理部に、第 2 の記憶制御装置が接続され、

前記第 1 の処理部が前記コマンドを受領した場合に、前記第 1 の記憶制御装置内の管理情報を参照して、前記コマンドが、前記第 1 の記憶制御装置が制御する記憶装置上の論理デバイスに対するものか、前記第 2 の記憶制御装置が制御する記憶装置上の論理デバイスに対するものを、判定することを特徴とする記憶システム。

【請求項 7】

請求項 6 に記載の記憶システムにおいて、前記判定が前記第 2 の記憶制御装置に上の論理デバイスに対するものであると判定した場合に、

前記第 1 の処理部または、前記第 2 の処理部が、前記コマンドを前記第 2 の記憶制御装置で制御可能なコマンドにコマンドの等価変換を行うことを特徴とする記憶システム。

処理部を介して前記第 2 の記憶制御装置に対して発行するかどうかを判定することを特徴とする記憶システム。

【請求項 8】

上位装置から受領した記憶装置に対するデータの読み出しを制御する第 1 の記

憶制御装置を有する記憶システムであって、

前記第 1 の記憶制御装置は、

データを蓄積する第 1 の記憶装置と、

前記上位装置に接続され、前記上位装置からの読み出しコマンドを処理する第 1 の処理部と、

前記第 1 の処理部の処理結果に応答して前記第 1 の記憶装置から前記データを読み出して、キャッシュメモリに保持する第 2 の処理部と、を有し、

前記記憶システムは、前記第 2 の処理部と接続される第 2 の記憶制御装置を有し、

前記第 2 の記憶制御装置は、

前記第 2 の処理部からの読み出しコマンドを受領する第 3 の処理部と、

前記第 2 の記憶制御装置が制御する第 2 の記憶装置とを有し、

前記第 3 の処理部の処理結果に応答して前記第 2 の記憶装置からデータの読み出しを行う第 4 の処理部とを有することを特徴とする記憶システム。

【請求項 9】

請求項 8 に記載の記憶システムにおいて、

前記第 1 の記憶制御システムは、前記第 2 の記憶装置上に生成された論理デバイスの管理情報を保持するメモリを有することを特徴とする記憶システム。

【請求項 10】

請求項 9 に記載の記憶システムにおいて、

前記管理情報が、前記第 2 の記憶装置上の論理デバイスを含むペア論理デバイスの管理情報であることを特徴とする記憶システム。

【請求項 11】

請求項 10 に記載の記憶システムにおいて、

前記管理情報が、前記第 2 の記憶装置上の論理デバイスを含むペア論理デバイス間の差分情報を含むことを特徴とする記憶システム。

【請求項 12】

請求項 11 に記載の記憶システムにおいて、

前記第 1 または、前記第 2 の処理部が、前記差分情報に基づいて、前記上位装置

から読み出し命令を受領したデータを 前記第1の記憶制御装置が制御する論理デバイスから読み出すか、前記第2の記憶制御装置が制御する論理デバイスから読み出すかを判定することを特徴とする記憶システム。

【請求項13】

第1の記憶装置を有する第1の記憶制御装置に、第2の記憶装置を有する第2の記憶制御装置が接続された記憶システムの制御方法であって、

前記第1の記憶装置が、上位装置からデータの書き込みまたは読み出し命令を受領する第1のステップと、

前記第1の記憶制御装置内の管理情報に基づいて、前記上位装置からのデータの書き込みまたは読み出しを前記第1の記憶装置に対して行うか、前記第2の記憶装置に対して行うかを判定する第2のステップと、

前記第2のステップの判定結果に応答して、前記第1または前記第2の記憶装置に対してデータの書き込みまたは読み出しを行う第3のステップとからなる記憶システムの制御方法。

【請求項14】

上位装置から受領した記憶装置に対するデータの書き込みや読み出しを制御する第1の記憶制御装置を有する記憶システムにおいて、

前記第1の記憶制御装置は、前記上位装置に接続され、前記上位装置からのコマンドを処理する第1の処理部と、 前記上位装置から書き込み命令を受領したデータを記憶する第1の記憶装置と、前記上位装置から書き込み命令を受領したデータまたは前記上位装置から読み出し命令を受領し前記第1の記憶装置から読み出したデータを一時的に保持するキャッシュメモリと、 前記記憶システムの管理情報を保持するメモリと、 前記キャッシュメモリ上のデータの前記第1の記憶装置に対する書き込みを制御する第2の処理部と、を含み、 前記第2の処理部に、第2の記憶制御装置が接続される記憶システムの制御方法であって、

前記第1の処理部が、上位装置からデータの書き込みまたは読み出し命令を受領する第1のステップと、

前記上位装置から受領したコマンドが、前記第1の記憶装置上の論デバイスに関するものか、前記第2の記憶装置上の論理デバイスに対するものかを前記管理

情報を参照して判定する第 2 のステップと、

前記第 2 のステップにおける判定結果が、前記第 2 の記憶装置上の論理デバイスに対するものであるとの判定結果である場合に、

前記第 2 の処理部を介して、前記第 2 の記憶制御装置に制御コマンドを中継することを特徴とする第 3 のステップとを有することを特徴とする記憶システムの制御方法。

【請求項 1 5】

請求項 1 4 に記載の記憶システムの制御方法において、

前記第 2 の記憶制御装置に制御コマンドを中継する場合に、前記第 1 の処理部または、前記第 2 の処理部が、前記中継する制御コマンドを前記第 2 の記憶制御装置が処理可能なコマンドに等価変換する第 4 のステップとをさらに含むことを特徴とする記憶システムの制御方法。

【請求項 1 6】

請求項 1 4 に記載の記憶システムにおいて、

前記記憶システムは、前記第 1 の記憶装置上の論理デバイスと、前記第 2 の記憶装置上の論理デバイスとの間に論理デバイスペアを生成するステップをさらに有し、前記論理ペアデバイスに対する前記上位装置からの書き込み命令を受領した場合に、前記論理デバイスの更新情報を差分情報として前記管理情報に蓄積することを特徴とする請求項 1 4 に記載の記憶システム。

【請求項 1 7】

請求項 1 4 に記載の記憶システムが、前記上位装置からのデータの読み出しコマンドを受領した場合に、

前記差分情報を参照して、前記読み出しを前記第 1 の記憶装置上の論理デバイスと、前記第 2 の記憶装置上の論理デバイスのいずれから読み出すかを判定するステップとをさらに有することを特徴とする記憶システムの制御方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、2以上の記憶制御装置間での論理デバイスペア構成方法、及び論理

デバイスペアへの各記憶制御装置からの I / O の処理方式に関する。

【 0 0 0 2 】

【従来の技術】

複数の記憶制御装置に跨って論理デバイスペアが構築されている場合、この論理デバイスペアをそれぞれ別にアクセス可能にしようとする場合、上位装置から Split 状態と呼ばれる状態に遷移するコマンドを発行し、ペアを構成されている論理デバイスのそれぞれは、アクセス履歴を管理される状態に遷移させられる。このようなリモートコピーに関する技術として例えば特許文献 1 がある。

【 0 0 0 3 】

この場合、上位装置からのそれぞれの論理デバイスに対する I / O 命令は、それぞれの記憶制御装置にが受領し、処理をされる。上位装置からの I / O をそれぞれの記憶制御装置で処理するために、論理デバイスペア間ではデータの差異が生じる。この生じたデータの差異は、それぞれの記憶制御装置上に持つ差分情報蓄積手段に記録される。これら独立して制御されていた論理デバイスをもとの同期しているペアの状態にするには、Resync というコマンドを上位装置から発行する。このリンシンクコマンドが発行されると、正論理デバイスと副論理デバイスのデータを一致させるために、正論理デバイスのデータを副論理デバイスにコピーする。ここでいう正論理デバイスとは、例えば、運用系の論理デバイスのことで、副論理デバイスはある時点での運用系の論理デバイスのデータイメージを保持しているボリュームを指す。

【 0 0 0 4 】

但し、この場合において、Split 中にデータを更新するような I / O を受けていないデータについては、コピー量・コピー時間の短縮の目的から、それぞれの記憶制御装置上で記録した前述の差分情報をマージし、差異部分についてのみコピー処理を行っている。例えば、Resync による差分データのコピー処理を、正側制御装置が主導で行う場合に、副側制御装置に記録された副側論理デバイスの差分情報を、副側制御装置に対し転送指示を与え、正側制御装置は受信した差分情報を正側論理デバイスの差分情報とマージして、差分コピーするデータを決定する。

【0005】

【特許文献1】 特開平10-333838号公報

【発明が解決しようとする課題】

前記従来技術によれば、正側記憶制御装置と副側記憶制御装置との間で、差分情報を送受信するためのインタフェースが不可欠である。現時点では、差分情報の管理方式は、ストレージメーカー独自の技術が採用されている。したがって、Resync時の差分コピー量・時間を短縮することの可能なリモートコピーの正副記憶制御装置は、同一メーカーのストレージ製品であることが要求される。。

【0006】

しかしながら、この制約は各社の記憶制御装置が存在するヘテロ環境の中では、記憶制御装置の構成の自由度を大きく妨げてしまう。

【0007】

そこで本発明の第一の目的は、リモートコピー時の正副制御装置間の差分管理方式に違いがあっても、高速なリシンクを可能とすることにある。また、副側制御装置に差分管理機能を持たない記憶制御装置上の論理デバイスとの間であっても、高速なリシンクを可能とすることにある。

【0008】

さらにまた、従来技術によれば、正側論理デバイスと副側論理デバイスの更新情報をそれぞれの制御装置側で差分管理することから、ベアスプリット (Split) 状態での上位装置からのリード命令は、リード命令先の論理デバイスから読み出すしかない。その読み出そうとするデータが、更新されたものである場合には、更新されたデータを読み出す必要があるからである。

【0009】

そこで、本発明の第二の目的は、上位装置のリード命令を正副論理デバイスから読み出させるように分散処理を可能とする記憶制御装置を提供することにある。

【0010】

【課題を解決するための手段】

上記第1の目的を達成する本願発明の記憶制御装置は、副側論理デバイスを正

側制御装置と接続し構成することにより、上位装置の副側論理デバイスに対する I/O 命令を正側制御装置で受領する手段と、上位装置の副側論理デバイスに対する I/O 命令を正側制御装置から副側装置に対し実行する手段と、ペア Split 状態において論理デバイスペアに対するそれぞれ論理デバイスデータの更新 I/O 位置を差分情報化し記録する手段と、ペア Resync 時に記録した差分情報から差分コピーする手段と、各論理デバイスに対する I/O の統計情報と論理デバイスペアの差分情報から上位装置のリード命令に対しデータの読み出し先論理デバイスを決定する手段とを備える、リモートコピーシステムで達成できる。

具体的には、記憶制御装置における論理デバイスペア間の正側と副側差分情報を、正側記憶制御装置に持つように構成する。この差分情報を利用することで正副論理デバイスのデータの一致部分が把握できることから、上位装置の副側論理デバイスに対するリード命令を正側論理デバイスで読み出すことが可能かどうか判断可能になり、同様に正側論理デバイスに対するリード命令を副側論理デバイスで読み出すことが可能かどうかを判断できるようになる。

【 0 0 1 1 】

【発明の実施の形態】

図 1 は、本実施例のシステム構成を示す。第 1 の記憶制御装置 2 は、この記憶制御装置 2 に命令を発行する情報処理装置である上位装置 1（以下単に上位装置と呼ぶ。）に接続され、この上位装置からのデータ出入力要求に応答して、第 1 の記憶装置 3 に対して、データの書き込みや読み出しを制御する。

この第 1 の記憶制御装置 2 は、Initiator I/O 処理部を介して、第二の記憶制御装置 4 と接続される。

【 0 0 1 2 】

この第 1 の記憶制御装置は、上位装置 1 から書き込みや読み出し命令を受けたデータを一時的に蓄積するキャッシュメモリ 10 を有する。

【 0 0 1 3 】

このキャッシュメモリ 10 は、第 2 の記憶制御装置 4 に接続される第 2 記憶装置 5 a や、第 3 記憶装置 5 b から第 1 の記憶制御装置 2 が読み出したデータをバッファリングする場合や、第 1 の記憶制御装置 2 から、I/O 処理部 8 を介して

、第2の記憶制御装置4に接続される第2記憶装置5aや第3の記憶装置5bに書き込みするデータを一時的に保持するためにも用いられる。

【0014】

この第1の記憶制御装置2は、上位装置1からのI/O命令を処理するためのTargetコマンド処理部6と、第一記憶装置に対しI/O命令を処理するためのInitiator I/O処理部7とを備え、同様に第二制御装置に対しI/O命令を処理するためのInitiator I/O処理部8を備える。

【0015】

さらにこの第1の記憶制御装置2は、Targetコマンド処理部6とInitiator I/O処理部7、8との間で、情報の共有や交換を行うための共有メモリ9から構成される。共有メモリ9は、Initiator I/O処理部8を介して接続される第2の記憶制御装置4のアドレスを管理する制御装置情報11を有し、この情報によって、第1の記憶制御装置は、第2の記憶制御装置のアドレスを特定することが可能となる。さらに共有メモリ内には、論理デバイスと記憶装置と、記憶制御装置と、論理デバイスペアとを関連付けるためのデバイス情報12を有し、第一の記憶制御装置2は、この情報を参照して、上位装置が論理ボリュームに対して発行した命令が、第1記憶装置3のどのアドレスへのアクセス命令なのか特定を容易にする。

【0016】

一方、このデバイス情報12は、どの論理ボリューム同士がペアを構成しているかを管理する情報であり、第1の記憶制御装置2は上位装置1からのI/O要求に対して、この情報を参照して、ペアを形成している論理ボリュームが存在する場合には、当該ペアとなっている論理ボリュームにもデータの書き込みを行うような処理を実行する。即ち、これらの情報によって、書き込みを行うべき論理デバイスが自筐体に存在するか。第2の記憶制御装置4のように第1の記憶制御装置2とは別の筐体なのかを識別する。

【0017】

さらに、この共有メモリ9は、ペアを形成している論理デバイス間（以下論理デバイスペア間）でデータに差分が生じた場合に差分が生じた位置を示す差分情

報 13 を有する。

【0018】

一方、この共有メモリ 9 は、Target コマンド処理部 6 と Initiator I/O 処理部 7、8 との間で論理デバイスに対する処理の指示と処理の結果を連絡するための I/O 情報 14、及び、論理デバイスに対する I/O 処理数を記録するための統計情報 15 を備える。この統計情報 15 は、論理デバイスの物理ボリューム上の配置を最適化する場合に用いられる他、「I/O 命令を出すときの最適化」すなわち I/O のバランシングを目指す目的で使用する。

【0019】

第 2 の記憶制御装置 4 は、第 1 の記憶制御装置 2 の Initiator I/O 処理部 8 のから出力される、第 2 の記憶装置 5 a 上に配置される論理デバイス 17 と第 3 の記憶装置 5 b にある論理デバイス 19 に対する I/O 命令の処理手段を備える。具体的な構成は、第 1 の記憶制御装置 2 と同じである。

【0020】

図 2 は、記憶制御装置が、記憶装置上の論理デバイス配置を管理するための制御装置情報 11 を示す。この制御情報は、Initiator I/O 処理部に接続される制御装置を特定し、管理するための情報で、図 2 に示すようにテーブル構成を有する。

【0021】

記憶制御装置番号 101 は、第一の記憶制御装置 2 内で、Initiator I/O 処理部 8 に接続された記憶制御装置を識別するための識別子である。筐体位置 102 は、アドレスを管理する記憶制御装置が第一記憶制御装置自身を指すのか、第一の記憶制御装置の Initiator I/O 処理部 8 を通じて接続される他筐体であるのかを示すデータである。

【0022】

さらにこのテーブルは、制御装置が Initiator I/O 処理部を通じて接続される場合に、どの Initiator I/O 処理部を経由するかを示すイニシエータ制御部番号と、その制御装置がユニークに判別できるためのノードネーム 104 と、Initiator I/O 制御部がその制御装置に I/O を発行するための制御装置アドレ

ス105を情報として持つ。

【0023】

この情報は前述の通り、共有メモリ9内に置かれ、ターゲットとする論理デバイスを配下に有する記憶制御装置のアドレスを求めるときや、記憶制御装置間のコネクションを行うときに参照する。

【0024】

次に、デバイス情報12の構成要素を図3で説明する。上位装置がI/Oユニットとして認識する論理デバイス番号を配列として、参照する論理デバイスに関する情報の格納位置が求まる。制御装置番号202は、論理デバイスの実装される記憶制御装置の番号であり、図2の制御装置情報テーブルで管理される制御装置の番号（制御装置情報テーブルの配列の要素番号に等しい）である。、記憶装置番号203は、制御装置番号202が指す記憶制御装置内の記憶装置の番号である。

【0025】

ペア状態204は、論理デバイスがリモートコピーペアデバイスであることを示す情報で、図でSplitとあるのは、ペアを組んでいる論理デバイスが存在し、その論理デバイスペアが、現在それぞれ独立して書き込みや読み出しの要求を受け、その差異は、差分管理によって管理されている状態を示している。またSimplexとあるのは、どの論理デバイスとペアを構成しているほかの論理デバイスがないことを示している。

【0026】

ペア番号205は、そのペアを識別するための識別子である。属性206は、ペアの正側論理デバイスなのか副側論理デバイスなのか属性を表すものである。ペア論理デバイス番号207は、ペアを組む相手先の論理デバイス番号を示す。コピーポインタ208は、ペア形成中の初期コピー進捗位置を指すポインタである。

【0027】

デバイス情報12は、論理デバイスに対するアクセス要求に対し、論理デバイスの実装先、つまりデータをリード・ライトするためのI/Oを発行する物理的

な位置情報である記憶装置番号または記憶制御装置番号を求めるために使用する。また、ペア状態204から、データ更新I/Oについては、ペアとなる正副両論理デバイスに対する二重書き込みの要否を判断する。二重書き込みが必要と判断する場合には、ペアを組む相手先論理デバイス番号207を用いて、相手先論理デバイスの実装先をデバイス情報テーブル12から求める事が出来る。

【0028】

次に、I/O情報テーブルを図4を用いて説明する。Targetコマンド処理部6が上位装置から論理デバイスデータのアクセス要求を受け、キャッシュメモリ10と記憶装置3や第二の記憶制御装置4とでデータ転送の必要性が発生するとき、Initiator I/O処理部7、8に対しリード・ライトをリクエストするために、リクエストキュー301～307を使用する。リクエストキューポインタテーブル301は、リクエストキューに要求を格納する先を示すINポインタ302と、リクエストキューの要求を取り出すためのOUTポインタ303から成る。Initiator I/O処理部7、8は、論理デバイスに対する要求が発生しているかを知るために、INポインタ302とOUTポインタ303の差分が有るかを判断する。リクエストキューは、リクエストキューポインタのポイント先としてリクエストキューエントリ番号304と、要求する命令を格納するコマンド305、コマンド対象ブロック番号306、コマンド処理ブロック長307から成る。レスポンスキュー308～313は、リクエストキューで指示された要求の処理結果を格納するために使用する。

レスポンスキューポインタテーブル308は、レスポンスキューに応答を格納する先を指すINポインタ309と、応答を取り出す先を指すOUTポインタ310から成る。指示された要求に対する処理結果は、INポインタ309が指すレスポンスエントリ番号311の処理結果312に格納し、処理結果の詳細情報が有れば詳細情報313に格納する。

【0029】

換言すれば、I/O情報テーブルとは、論理デバイスに対するアクセス要求に対し、Initiator I/O処理部を通じて、物理デバイス（＝記憶装置）に対しI/Oを発行すべきと判断した場合に、Targetコマンド処理部6と、Initiator I/O処理部

7, 8間で、要求を通知し（リクエストキュー）、要求に対し応答を通知する（レスポンスキュー）するためのテーブルである。したがって、記憶装置3だけでなく、第二の記憶制御装置4とのデータ転送の必要性が発生した場合に使用する。

【0030】

次に、図5の差分情報13について説明する。Targetコマンド処理部6が、上位装置1から論理デバイスのデータ更新要求を受けたときに、デバイス情報12から論理デバイス番号を配列として、この論理デバイスのペア状態204を参照する。ペア状態がSplitにあるとき、ペアデバイスのどちらか一方だけがデータを更新することになる。このため、正副論理デバイス間で更新した部分は正副不一致になる。よって、更新位置と長さから不一致になるデータに相当する差分ビット位置を求め差分ビットを「1」に設定する。差分ビットの意味は、正副論理デバイス間でデータが一致となる位置の差分ビットは「0」、不一致となる位置の差分ビットは「1」としておく。このことから、ペアResyncするときは、そのペアの差分情報13の差分ビットが「1」の部分だけを、正側の論理デバイスから副側の論理デバイスに対し差分コピーすれば良い。

【0031】

次に、図6の統計情報テーブル15について説明する。上位装置1からの所定の論理デバイスデータに対するアクセス要求に対し、Targetコマンド処理部6がI/O情報テーブル14を用いて、Initiator I/O処理部7, 8を介して、記憶装置3, 5a, 5bにI/Oを指示したときに、統計情報テーブル15の対象論理デバイスの全I/Oカウンタ402のカウンタを更新する。そのI/Oがリード要求であった場合リードI/Oカウンタ403を、また、そのI/Oがライト要求であった場合ライトI/Oカウンタ404を更新する。この統計情報テーブル15により、上位装置1の論理デバイスに対する要求によって記憶装置に対して何回のI/Oを指示したかの回数を把握できる。

さらに、図7、図8を用いて、これらの制御情報を使用する本発明の方式について詳細に説明する。

図7, 8は、上位装置1から第1記憶制御装置2に対し発行される、第2の記憶

制御装置上の論理デバイス 5 a, 5 b に対するの I/O 要求を、第 2 記憶制御装置に I/O コマンドを転送するまでの流れについて説明している。

【0032】

図 7 は、上位装置 1 と接続される第 1 の記憶制御装置 2 の Target コマンド処理部 6 における処理フローである。Target コマンド処理部 6 は、受信した I/O 要求が自筐体内で処理可能か、それとも Initiator I/O 処理部 8 を介して他の記憶制御装置に転送するべきかを判断する。

【0033】

ステップ 1002 は、上位装置 1 から与えられた論理デバイス番号から、デバイス情報 12 の制御装置番号 202 を取得している。具体的には図 3 のテーブルを用いて、指定された論理デバイス番号 201 に対応する制御装置番号 201 を読み出す。

【0034】

続くステップ 1003 〜は、制御装置番号 202 から図 2 に示す制御装置情報 11 を用いて、筐体位置 102 を参照し、論理デバイスが自身が制御する記憶装置上の論理デバイスであるかどうかに関する情取得している。

続くステップ 1004 では、上位装置 1 からの I/O 要求が自筐体内でのみ処理可能かを判定する。自筐体内でのみ処理可能であるならば、通常の Target I/O 動作が可能であり、ステップ 1011 に進む。ここで、自筐体内でのみ I/O 動作が可能という意味は、上位装置 1 から書込みや読み込み等の I/O コマンドを受領した記憶制御装置自身の制御下に、その I/O コマンドの対象となる論理デバイスが、存在しているという意味である。

【0035】

自筐体で処理不可もしくは他筐体でも処理すべきと判断した場合、即ち、I/O の対象となった論理デバイスが、他記憶制御装置の制御下にある場合は、続くステップ 1005, 1006 にて、図 4 に示す I/O 情報 14 にある I/O メッセージテーブルのリクエストキューに空きエントリがあるかチェックする。リクエストキューに空きが無い場合は、ステップ 1012 に進み、エラーを返却することでリトライを催促する。リクエストキューにエントリ可能であれば、ステッ

ステップ1007でリクエストキューのINポインタ302が示すリクエスト番号のリクエストエントリ304に、上位装置1から受領した要求をコマンド305・ブロック番号306・ブロック長307を登録する。ステップ1008では、リクエストをリクエストキューに格納済みであることを示すために、リクエストキューのINポインタ302を加算する。ステップ1009では、リクエストキューに要求を転送したことを呼び元に通知するための返却値を設定する。

ところで、図7においては、上位装置1からのI/Oコマンドがアクセスの対象としている論理デバイスが、当該I/Oコマンドを受領した記憶制御装置自身が制御する記憶装置上のデバイスであるかどうかを判定し、I/Oコマンドを受領した記憶制御装置自身の論理デバイスを対象とするアクセスである場合には、その記憶制御装置自身が、自身の制御下の論理デバイスに対するI/O処理を行ない、そうでない場合は、他の記憶制御装置の筐体番号を特定して、当該他の筐体の記憶制御装置にその処理を委ねるように構成した場合を説明した。この点を図1を用いて詳述すると、上位装置1から受領したI/Oコマンドが、第1の記憶制御装置2の制御下にある第1記憶装置3上に形成された論理デバイス18である場合は、Targetコマンド処理部6及びInitiator I/O処理部7が共有メモリ9内の各種情報に基づいて、論理デバイス18へのI/O処理を行う。一方、上位装置1から受領したI/Oコマンドが、第2記憶装置5a、または、第3記憶装置5b上の論理デバイス17、または、19を対象としたものである場合は、Targetコマンド制御部6は、Initiator制御部8を介して、当該I/Oを第2記憶制御装置4に中継するのである。

【0036】

I/Oコマンドの中継を受領した第2記憶制御装置4は、当該I/Oコマンドが第2記憶制御装置自身が制御する記憶装置上の論理デバイスを対象としたものかどうかを判定する。(図7の手順に同じ。)そして、このI/Oコマンドが対象とした論理デバイスが、例えば第2記憶装置5a上の論理デバイス17である場合には、当該論理デバイス17に対して第2記憶制御装置内の処理部がI/O処理を行うのである。

【0037】

ここで、第 1 の記憶制御装置 2 と、第 2 の記憶制御装置 4 とは同じ種類または性能の記憶制御装置である必要はない。なぜなら、第 2 の記憶制御装置は、第 1 の記憶制御装置 2 から中継される I/O コマンドを単に、上位装置から受領する I/O コマンドであるのと同じように受領し、処理するからである。

【 0 0 3 8 】

一方、第 1 記憶制御装置 1 の Initiator I/O 処理部を介して中継される I/O コマンドは、必ずしも上位装置 1 が発行した I/O コマンドをそのまま、転送するものでなく、第 1 記憶制御装置 2 の Target コマンド処理部 6 或いは、Initiator I/O 処理部 8 が何らかの変換を行ってから中継するように構成することもできる。

【 0 0 3 9 】

上位装置 1 から発行されるコマンドが、第 2 の記憶制御装置 4 自身では処理できないような I/O コマンドであったとしても、第 1 の記憶制御装置が当該 I/O コマンドを解釈して、第 2 の記憶制御装置 4 が処理できる等価の I/O コマンドとして発行するのである。

【 0 0 4 0 】

このようにコマンドを変換してから中継することが可能に構成にすることは、性能の異なる記憶制御装置を利用する場合に最適である。

【 0 0 4 1 】

例えば、第 2 の記憶制御装置を介して、その制御下の第 2 記憶装置 5 a、第 3 記憶装置 5 b を利用していたユーザが、新たに新しい機能を有する記憶制御装置として、第 1 の記憶制御装置 2 を導入したような場合である。

【 0 0 4 2 】

本願発明を適用しない場合、新たに便利な I/O コマンドを有する第 1 の記憶制御装置 2 を導入したとしても、第 2 の記憶制御装置 4 は、従来の制御装置なので、ユーザがこれまで利用していた第 2 記憶装置 5 a や第 3 記憶装置 5 b 内のデータに対してはその新たな I/O コマンドを利用することができない。そのようなコマンドを利用するためには、第 2 記憶装置 5 a 及び第 3 記憶装置 5 b 内のデータを新たに導入した第 1 記憶制御装置の制御下の第 1 記憶装置 3 に移行しなけ

れなければならない。当然移行してしまうと、従来利用していた第 2 記憶制御装置 4 及び、その制御下の第 2 記憶装置 5 a 及び第 3 記憶装置が無駄になってしまう。

【 0 0 4 3 】

そこで、本願発明のように第 1 の記憶制御装置 2（新装置）で上位装置からの高度なコマンドを受領し、そのコマンドの解釈して、第 2 の記憶制御装置が処理できるコマンドするコマンド等価変換をこの第 1 の記憶制御装置 2 で行うと共に、第 1 の記憶制御装置 2 の共有メモリ 9 に蓄積された制御装置情報 1 1 を用いて、第 2 記憶制御装置 4（旧装置）の論理デバイスへのアクセスを可能にすれば、これまでのデータ資源及び、第 2 の記憶制御装置 4、記憶装置 5 a、5 b のようなハードウェア資源も有効に利用することができるのである。

【 0 0 4 4 】

ところで、本願図 1 においては、説明を容易にするために、上位装置と接続するアダプタであって、上位装置からのコマンドを受領して処理する Target コマンド処理部 6 が一つの場合を用いて説明したが、この上位装置と接続され、当該上位装置のコマンドを処理するアダプタ部は、通常 1 の記憶制御装置に対して、複数設けられることが一般的である。

【 0 0 4 5 】

そして、記憶システムの全体の拡張性、スケラビリティを向上させるために、できるだけ多くこのようなアダプタ部を設けるように工夫されている。このような上位装置と接続されるアダプタ部は、ディレクター部或いは、チャネルアダプタと呼ばれることがある

本願発明においては、第 2 の記憶制御装置 4 を第 1 の記憶制御装置 2 と接続するにあたり、このような TARGET コマンド部 6 を使用せずに、記憶装置を接続するアダプタである Initiator I / O 処理部を利用する。この 1 実施態様においては、その第 2 の記憶制御装置 4 を第 1 の記憶制御装置 2 が、記憶装置を接続するアダプタポートを介して接続されるので、第 2 の記憶制御装置 4 への接続のために、TARGET コマンド部 6 のような上位装置との接続部が利用されない。つまり、上位装置との接続アダプタを、他記憶制御装置へのコマンド中継によって占有されるという不便がないのである。

【0046】

図7のステップ1004は、単にホストがターゲットとした論理デバイスが自筐体に存在するかどうかということだけでなく、その処理を「自筐体では処理不可もしくは他筐体にも処理させるべき」という判断をするステップとすれば、上位装置からのコマンドをコマンドの種別に対応して、その処理を行う記憶制御装置を峻別することが可能となる。

【0047】

図8は、図7のステップ1007の手順で蓄積された上位装置1からのI/O要求に対し、Initiator I/O処理部8で要求を検出し、第二記憶制御装置4に対し要求を発行する手順を示す。

【0048】

図1に照らして具体的に説明すると、図7は、上位装置1から発行されたI/Oコマンドが、第1の記憶制御装置2内で処理されるべきか、Initiator I/O処理部8を介して、第2の記憶制御装置4内で処理されるべきものが判定する手順であり、図8は、第2の記憶制御装置4で処理すべき上位装置1からのI/OコマンドをInitiator I/O処理部が、処理する手順を示している。

【0049】

Initiator I/O処理部は、Targetコマンド処理部からコマンドを受領すると、自身がホストコンピュータのように振る舞い、第2の記憶制御装置4に対して、受領したコマンドを発行する。

【0050】

ステップ2002～2003は、initiator I/O処理部8が、自身に接続される記憶制御装置が制御する記憶装置上の論理デバイスに対する要求を蓄積する領域を監視して、その論理デバイスに対するI/O要求があるかどうかを判定するステップである。

Initiator I/O処理部は、自身に接続される記憶制御装置の制御下に生成された論理デバイスに対するリクエストがあるかどうかを、図4に示すI/O情報14のリクエストキューを参照して、判定する。

【0051】

具体的には、Initiator I/O処理部は、図4の該当する論理デバイスのI/OメッセージテーブルのリクエストキューINポインタ302とOUTポインタ303とが等しければ、INポインタ302が進んでいないので要求が無いと判断し、ステップ2011に進む。もしステップ2003で、INポインタ302とOUTポインタ303とが不一致であれば、リクエストが蓄積されていると判断して、続くステップ2004でリクエストキューOUTポインタ303が示すリクエスト番号のリクエストエントリ304から、コマンド305・ブロック番号306・ブロック長307を取得する。

【0052】

続くステップ2005は、リクエストを受けた論理デバイスの論理デバイス番号から、デバイス情報12を参照して、記憶制御装置及び記憶装置番号とを取得する。

【0053】

それに続くステップ2006で、制御装置情報11から記憶制御装置のアドレスを取得して、ステップ2007で、Initiator I/O処理部から、第2の記憶制御装置に対してコマンド送信する。ステップ2008は、リクエストに対応したことを示すために、リクエストキューのOUTポインタ303を加算する。ステップ2009では、リクエストキューから他制御装置に対しコマンド送信できたことを返却するための返却値を設定する。

【0054】

コマンド送信を受領した記憶制御装置は、上位装置から受領したコマンドと同様に、当該コマンドに対応する処理を行う。

【0055】

Initiator I/O処理部は、ステップ2007でコマンドを送信する前に、コマンドの処理内容を解釈し、送信先となる記憶制御装置の性能で処理が可能なコマンドに変換してから送信するようにしても良い。このようにすることで、性能や機能が違う記憶制御装置を第2の記憶制御装置として接続することが可能となる。

【0056】

次に、ペアResync時のコピーデータが、正と副のペアデータの差分に限定されるようにするための差分設定処理を説明する。ペアデバイスは少なくとも2つの論理デバイスからなる。本願説明においては、便宜上、ペアの片方の論理デバイスを正論理デバイス、もう片方の論理デバイスを副論理デバイスとして説明する。ペアを生成した論理デバイス間には、次のような状態がある。

第1の状態は2重化状態であって、これは正の論理デバイスと同期して副の論理デバイスが同じデータイメージになるように更新される状態である。

第2の状態は、スプリット (split) 状態であって、この場合は、上述の二重化状態が解消され、各論理デバイスは別々にデータの書き込みや読み出しのアクセスを受ける状態である。この場合、書き込みを受けた領域は図5に示す差分情報で管理され、どの位置が更新されたかが判別可能に管理される。

第2の状態であるスプリット状態から、再度二重化状態に戻す場合は、この差分情報を参照して、正論理ボリュームと副論理ボリュームが、同じデータイメージとなるようにデータを復元する。この処理をリシンク (Resync) という。図1の本発明の構成と、図7、図8で説明した通り、上位装置1から受領する第二記憶装置4・第二記憶装置5a・第三記憶装置5b上の論理デバイスへのI/Oは、第一記憶制御装置2のTargetコマンド処理部6で受け付ける。

【0057】

図9は、Targetコマンド処理部6が更新I/Oを受け付けたときに、差分情報13にある差分ビットを設定するための処理である。ステップ3001～3007と図5に示す差分情報は、正副論理デバイス毎で分けて処理する必要は無く、差分管理は正副デバイスを一元管理している。

ところで、従来、物理的にことなる記憶制御装置間に生成されたでペア論理デバイスは、各々の記憶制御装置において、その差分管理をしていた。併しながら、本願発明の正論理デバイスと副論理デバイスは、図8に示すように第2の記憶制御装置4の制御下の記憶装置5a、5b上に配置される論理デバイスに対するI/Oは、上位装置1から第1の記憶制御装置2が受取り、解析して第2の記憶制御装置4に発行する中継処理を行うことから、正副デバイスペア間のデータ不一致に対する差分管理を、第二の記憶制御装置が行う必要が無く（管理情報を持た

なくても良い。) 、第 1 の記憶制御装置 2 が管理するだけで良い。つまり、第 2 の記憶制御装置が、ダイレクトに上位装置からの I / O 要求を受領せずに、第 1 の記憶制御装置を介して受領するので、第 1 の記憶制御装置でペア論理デバイスのそれぞれの更新状態を一元管理すれば良いということになる。また、この場合において論理デバイス毎に差分を管理する方法を用いても良いが、スプリット状態の終了後、正論理デバイスのデータイメージに復元する場合、それぞれの差分をマージすることになるので、最初から、個々の論理デバイス単位でなく、ペア単位で差分を管理するようにしておくことでマージという手順が必要なくなるように構成しても良い。

【 0 0 5 8 】

この第 1 の記憶制御装置 2 における差分情報の一元管理のさらに優れている点については後述する。

【 0 0 5 9 】

図 9 の 3 0 0 1 の差分設定判断の手順を詳述すると、まず、ステップ 3 0 0 2 で、図 3 に示すデバイス情報から、ペアの状態を取得する。

ペアの状態がスプリット状態でない場合は、差分情報を取得する必要がないので終了する。(ステップ 3 0 0 7)

一方、二重化状態である場合には、更新ブロック番号から差分ビット位置を算出し、(ステップ 3 0 0 4) 、さらに更新ブロック長から、差分ビット数を算出する。(ステップ 3 0 0 5) 。

ステップ 3 0 0 4 及びステップ 3 0 0 5 で得られた情報に基づいて図 5 に示す差分管理テーブルの該当個所のビットをオンにする。

【 0 0 6 0 】

図 1 0 は、ペアリシンク (Resync) 時の処理フローである。前述の通り、正副デバイス間の差分情報が、第 1 の記憶制御装置 2 に保持する差分情報 1 3 で一元化されるため、図 1 0 の処理フローに、第 2 の記憶制御装置 4 からの差分情報を読み込む処理を含む必要が無い。

【 0 0 6 1 】

図 1 0 の差分コピー判定の処理は、差分ビットテーブルを参照して、差分ビッ

ト位置 0 から、差分ビットの参照を開始し、(ステップ 4002、4003)、差分ビットが 1 である場合には、そのビット位置に対応する正論理デバイスの内容を副論理デバイスにコピーする。(ステップ 4005)

順次差分ビット位置を加算して(ステップ 4006)差分ビット位置が、終端となるまで繰り返す。(4007)

図 11、12 は、論理デバイスペアのデータに対し、上位装置 1 からリード要求が出されたときの、正副論理デバイスのどちらからデータをリードするかを決定するための処理を示している。

【0062】

前述のとおり、本願発明においては、第 1 の記憶制御装置 2 とは、別体である第 2 の記憶制御装置 4 の制御下の記憶装置上の論理デバイス 17、19 への上位装置からの I/O が、第 1 の記憶制御装置 2 を介して行われるので、第 1 の記憶制御装置上でそのような他の記憶制御装置上の論理デバイスの差分情報をも管理することができる。

【0063】

そこで本願発明においては、上位装置が発行するリード命令に対して、より効率の良い論理デバイスから読み出すという方法を提供することが可能となる。即ち、ペアとなっている論理デバイスが二重化状態 (Duplex) の場合のみならず、スプリット状態であった場合でもペア単位で管理している差分情報を参照すれば、差分ビットに 1 が設定されていない場合は、どちらから読み出しを行っても良いことになるので、記憶装置に対する I/O の負荷を分散することができる。例えば、図 1 との論理デバイス 17 と論理デバイス 19 のように物理的に異なる記憶装置上の論理デバイスにおいて I/O を分散処理することで特定の記憶装置に対する負荷が増大して、ホストに対するレスポンス性能が劣化することを防ぐのである。

【0064】

図 11 を用いて本願発明のリードするデバイスの決定方法を詳細に説明する。ステップ 5002 はリード要求を受領した直後の初期状態を示す。本説明において、入力論理デバイスとは、ホストから処理対象として入力を受けた論理デバイ

ス番号を指し、処理論理デバイスとは、実際のリード処理を行う対象とする論理デバイスをいう。

初期状態において、入力論理デバイスと処理論理デバイス番号は一致している。

【 0 0 6 5 】

ステップ 5 0 0 4 において、入力論理デバイスのペア形成状態を取得し、続くステップ 5 0 0 4 で、そのペア状態が、スプリット (Split) 状態または二重化状態 (Duplex) にあるのか、それ以外かを判定する。それ以外の状態としてシンプレックス (simplex) というペアが形成されていない状態がある。ここで、スプリット状態でも二重化状態でもない場合には、ペアが存在しないこと (シンプレックス) であることが分かるので、そもそも他の論理デバイスを用いて読み出すことができず、処理を終了する。この場合は初期値がそのまま維持され、入力論理デバイス番号は、処理論理デバイス番号に一致し、処理論理デバイス番号の論理デバイスに対して処理を行うことは、ホストからリード要求を受けた論理デバイスに対してリード処理がなされることとなる、

それに対して、入力論理デバイスが、スプリット状態または、二重化状態である場合には、I/O の負荷比較 (ステップ) 5 0 0 5 を実行する。このステップ 5 0 0 5 については、図 1 2 を用いて詳述する。

【 0 0 6 6 】

図 1 2 は、ペアとなっている正副の論理デバイスのどちらに、ホストからの I/O 負荷がかかっているかを、判定する一実施例である。

まず、入力論理デバイスの I/O カウンタを取得し (6 0 0 2)、次いで、その入力論理デバイスのペアとなっている論理デバイスの論理デバイス番号を取得し (6 0 0 3)、さらに、その論理デバイス番号で特定される論理ボリュームの I/O カウンタを取得する。(6 0 0 4)

ステップ 6 0 0 5 において、入力論理デバイスの I/O と、そのペアと成っている論理デバイスの I/O とを比較し、入力論理デバイスの I/O カウンタの方が、ペア論理デバイスの I/O カウンタよりも大きい場合は、戻り値として、ペア論理デバイス番号を取得する。

【 0 0 6 7 】

一方、入力論理デバイスの I / O カウンタが、ペア論理デバイスの I / O カウンタよりも大きくない場合は、戻り値として入力論理デバイス番号を取得する。

【 0 0 6 8 】

この 1 連のステップにおいては、I / O カウンタの少ない方の論理デバイス番号を取得することができる。

【 0 0 6 9 】

図 1 2 に示す処理を実行して、戻り値を取得すると、図 1 1 のステップ 5 0 0 6 において、処理論理デバイス番号にその図 1 2 の処理の戻り値である論理デバイス番号を設定する。

【 0 0 7 0 】

次に、リードブロック番号から、差分ビットの位置（ステップ 5 0 0 7）及びそのブロック長から差分ビット数を算出し（ステップ 5 0 0 8）、その領域の差分ビットテーブルを参照する。すべての差分ビットが 0 である場合は、どちらの論理デバイスから読み出しても良いことになるので、処理デバイスはステップ 5 0 0 5 で取得した I / O カウンタの少ない方の論理デバイスから読み出すのがふさわしい、そこで、処理論理デバイス番号として、ステップ 5 0 0 5 で取得した I / O カウンタの少ない方の論理デバイス番号が、処理論理デバイスとなる。

【 0 0 7 1 】

一方、ステップ 5 0 0 1 で差分ビットが 0 でない場合は、処理論理デバイス番号と入力論理デバイス番号とを再度一致させ、入力論理デバイスに対して読み出し処理を行うようにしている。これは、読み出そうとする領域が正・副の論理デバイス間で異なっていることを示しているので、読み出す論理デバイスを変更することができないからである。

【 0 0 7 2 】

図 1 1 及び図 1 2 に示す読み出し論理デバイスの決定手順は、その手順を適宜変更することができる。

例えば、図 1 1 のステップ 5 0 0 7 から 5 0 1 0 の処理である差分ビットの状態判定をステップ 5 0 0 5 の I / O 負荷比較処理に先行して行うようにしても良い。その手順によれば、差分ビットに 1 が設定されている場合には、入力論理デバ

イスに対してリード処理を行う他ないのであるから、I/O負荷比較を行う必要がなくなる。

【 0 0 7 3 】

本願発明を換言すれば、異なる記憶制御装置にある論理デバイスとの差分情報を、第1の記憶制御装置で一元管理しているために、リード要求の出された位置のデータが正副論理デバイスで一致しているかどうかを判断できるため、Targetコマンド処理部6で、リード先の論理デバイスを選択することが可能になるということである。ま

尚、本願実施例においては、I/Oカウンタ値によって、処理論理デバイスを選択するように構成した例を示したが、システム全体の時間的な負荷バランスに特定の周期性があるような場合には、所定のタイムスケジュールによって、処理論理デバイスの選択を行うように設計することも可能である。

図13に本願の他の実施例を示す。

図13は、ディスクドライブ（第一の記憶制御装置）を接続するためのポートにファイバチャネルデバイスコネクタを介在させて、第2の記憶制御装置とを接続したものである。

この場合、第1の記憶制御装置2の共有メモリ上のデバイス情報テーブルには、ファイバチャネルデバイスコネクタのどのノードに第2の記憶制御装置が接続されているかについての情報を保持する必要がある。

【 0 0 7 4 】

図14は、図13に示す第2の記憶制御装置の接続するノードをノードネーム104として、デバイス情報テーブルに保持した一例を示している。

図13をさらに詳細に説明すれば、図13は、Targetコマンド処理部6に加えて、さらにTargetコマンド処理部70を有する構成である。また、Initiator I/O処理部7を介して、

ファイバチャネルデバイスコネクタ90が接続され、このコネクタ90bに第1の記憶装置が接続され、他のコネクタ90cに第2の記憶制御装置4が接続されている。

【 0 0 7 5 】

このような構成によって、Initiator I / O 処理部 7 から第二の記憶制御装置 4 配下の第二、第三の記憶装置 5 a、5 b を制御可能な構成としている。

図 1 4 は、図 1 3 で第 2 の記憶制御装置 4 が、Initiator I / O 処理部 7 で接続されていることで、他筐体である第 2 の記憶制御装置 4 との接続口を示すイニシエータ制御部番号 1 0 3 が、自筐体接続口のイニシエータ制御部番号 1 0 3 と同じ値を指すことを示している。

図 1 5 さらに他の実施例を示す。

図 1 5 は、複数の記憶制御装置が、Initial I / O 処理部を介して、複数を接続されている状態を示している。本発明におけるデバイス情報や制御手段を有する複数の記憶制御装置 2、4 x、4 y によりを経由してアクセス可能な記憶制御装置 4 z の制御下に実装される記憶装置 5 z にデータが格納される論理デバイス 1 9 z と、記憶制御装置 2 の制御下に実装される記憶装置 3 にデータが格納される論理デバイス 1 8 とが、論理デバイスペアを構成している図を示す。このように、上位装置 1 から論理デバイス 1 9 z に対するアクセス要求が発生した場合、記憶制御装置 2、4 x、4 y の共有メモリに持つ制御装置情報・デバイス情報・I / O 情報を用いて、論理デバイスの特定を行い、異なる記憶制御装置に対して、処理要求に対する処理が可能となる。

【 0 0 7 6 】

本願発明は、図 1 で示すとおり、第 1 の記憶制御装置 2 の Initiator I / O 処理部に、別の筐体で第 2 の記憶制御装置 4 を接続した構成である。換言すれば、第 1 の記憶制御装置のディスクをつなぐ口を利用して、他の記憶制御装置 4 と接続し、第 1 の記憶制御装置 2 の制御下で、当該第 2 の記憶制御装置を制御しようとするものである。

【 0 0 7 7 】

本願発明のようなディスクアレイ装置の場合、ディスクアダプタは複数のディスクドライブをファイバチャネル接続することから、複数のディスクドライブをつなぐ装着口を持っていて、その装着口はディスクドライブをフル実装しない限り使いきらない。ディスクドライブの装着口を、本発明にある第 2 の記憶制御装置 4 との接続に用いる事で、第一の記憶制御装置 2 と上位装置 1 とのチャネル接

続口数が制限されることもない。

【 0 0 7 8 】

第 1 の記憶制御装置 2 に接続する記憶装置の代わりに、第 2 の記憶制御装置 4 を接続し、その第 2 の記憶制御装置 4 が見せる論理デバイスが、第 1 の記憶制御装置 2 の記憶装置として制御することが可能になる。

【 0 0 7 9 】

本願発明の図 1 においては、第 1 の記憶制御装置 2 のディスク接続ポートに接続される第 2 の記憶制御装置が 1 台である場合を例示して詳細に説明したが、何台の記憶制御装置が、第 1 の記憶制御装置の配下に接続されるかは随時、ユーザの設定ポリシーに従って設計可能である。

【 0 0 8 0 】

また、図 1 5 に示すように第 2 の記憶制御装置のディスク接続ポートに他の記憶制御装置をさらに接続して、階層的に制御することも可能である。

【 0 0 8 1 】

尚、本願発明において、Target コマンド処理部として説明した部位は、記憶制御装置のチャネアダプタとしての機能を有し、上位装置からの受領したコマンドを処理するように構成し、Initiator I / O 処理部は、記憶制御装置のディスクアダプタとしての機能を有し、記憶装置へのキャッシュメモリ上のデータの書き込みも制御するに構成することができる。

【 0 0 8 2 】

【発明の効果】

本発明により、第 2 の記憶制御装置の配下にある論理ボリュームの管理が、第 1 の記憶制御装置によって一元管理されることが可能になる。また、この場合において、第 2 の記憶制御装置が有する記憶装置の接続ポートを利用すれば、第 1 の記憶制御装置の上位装置側に接続されるポートが、リモートコピーペアとのインタフェースとして利用されることがなくなり、接続可能な上位装置台数が減ることがない。

【 0 0 8 3 】

さらに、物理的に異なる筐体間の論理デバイスペアの副側制御装置の差分管理

方式の違いや、その有無に関わらず、ペアを生成することが可能となり、Resync時間の短縮も可能になる。

【 0 0 8 4 】

また、正側記憶制御装置上に、正と副両方の論理デバイス間の差分を保持するため、上位装置からのリード命令を、正副論理デバイスで負荷分散することが可能になる。

【図面の簡単な説明】

【図 1】 本発明のシステム構成である。

【図 2】 制御装置情報である。

【図 3】 デバイス制御情報である。

【図 4】 I/O要求・応答通知用のメッセージテーブルである。

【図 5】 差分情報である。

【図 6】 統計情報である。

【図 7】 I/O要求送信処理フローである。

【図 8】 I/O要求受信処理フローである。

【図 9】 差分ビット設定判定処理フローである。

【図 10】 差分コピー処理フローである。

【図 11】 正副リードデバイス決定処理フローである。

【図 12】 正副デバイス負荷判定フローである。

【図 13】 ディスクドライブ接続口に第二の記憶制御装置をつないだ構成図である。

【図 14】 図 13における制御装置情報である。

【図 15】 第2の記憶制御装置のディスク接続ポートに他の記憶制御装置をさらに接続した図である。

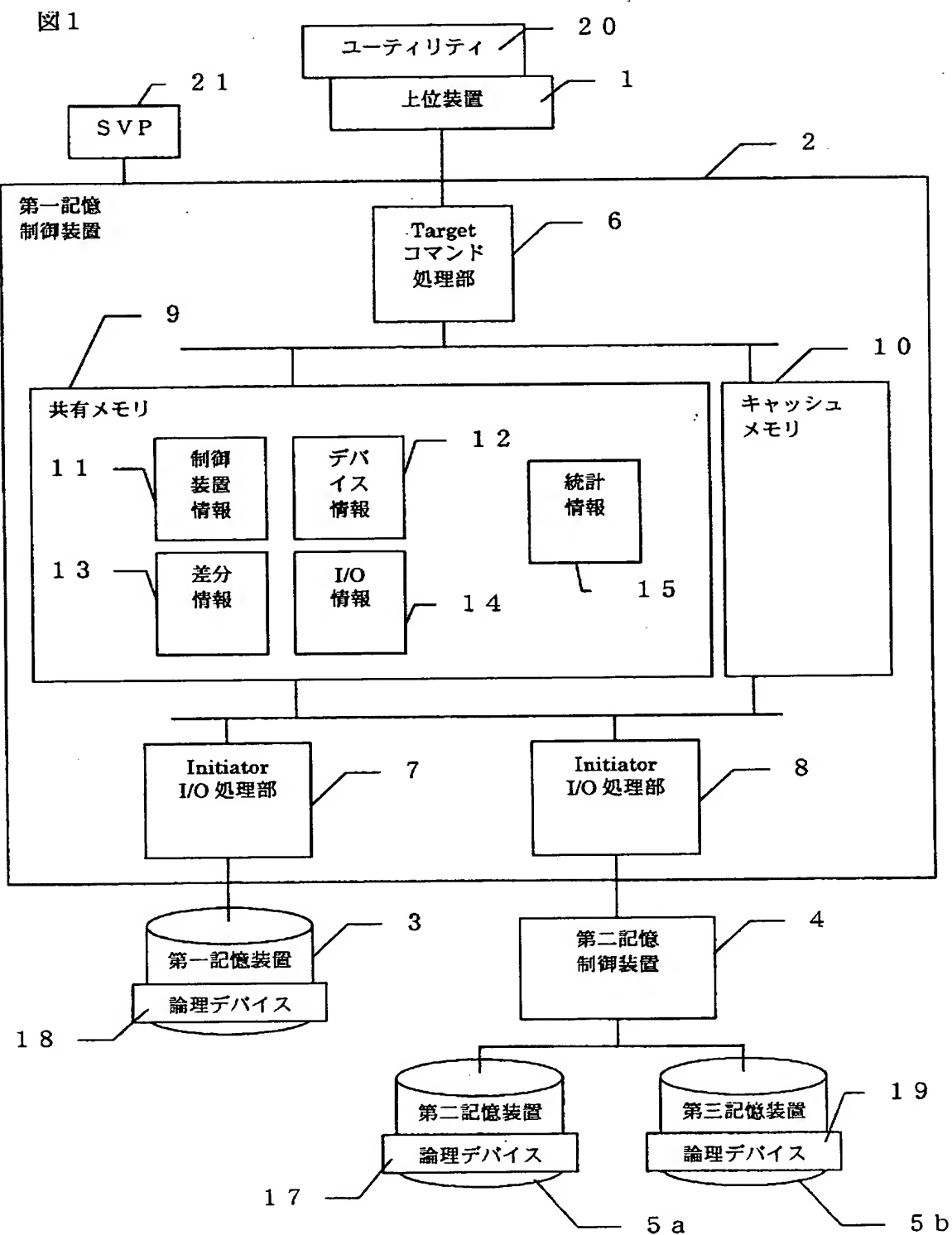
【符号の説明】

1・・・上位装置、2・・・第1記憶制御装置、3・・・第記憶装置、4・・・第2記憶制御装置、5 a・・・第2記憶装置、5 b・・・第3記憶装置、6・・・TARGETコマンド処理部、7・・・Initiator I/O処理部、9・・・共有メモリ、10・・・キャッシュメモリ、11・・・制御装置情報、12・・・デバイス情報、13・・・差分情報、14・・・I/O

○情報、 1 5 ……統計情報、 1 7 ・ 1 8 ・ 1 9 ……論理デバイス

【書類名】 図面

【図 1】



【図 2】

図 2

101 制御装置 番号	102 筐体 位置	103 イニシエータ 制御部番号	104 ノード ネーム	105 アドレス
0	自	0	HTC	0000E8
1	他	1	SBC	0100E4
2	無効	無効	無効	無効
.
.
.

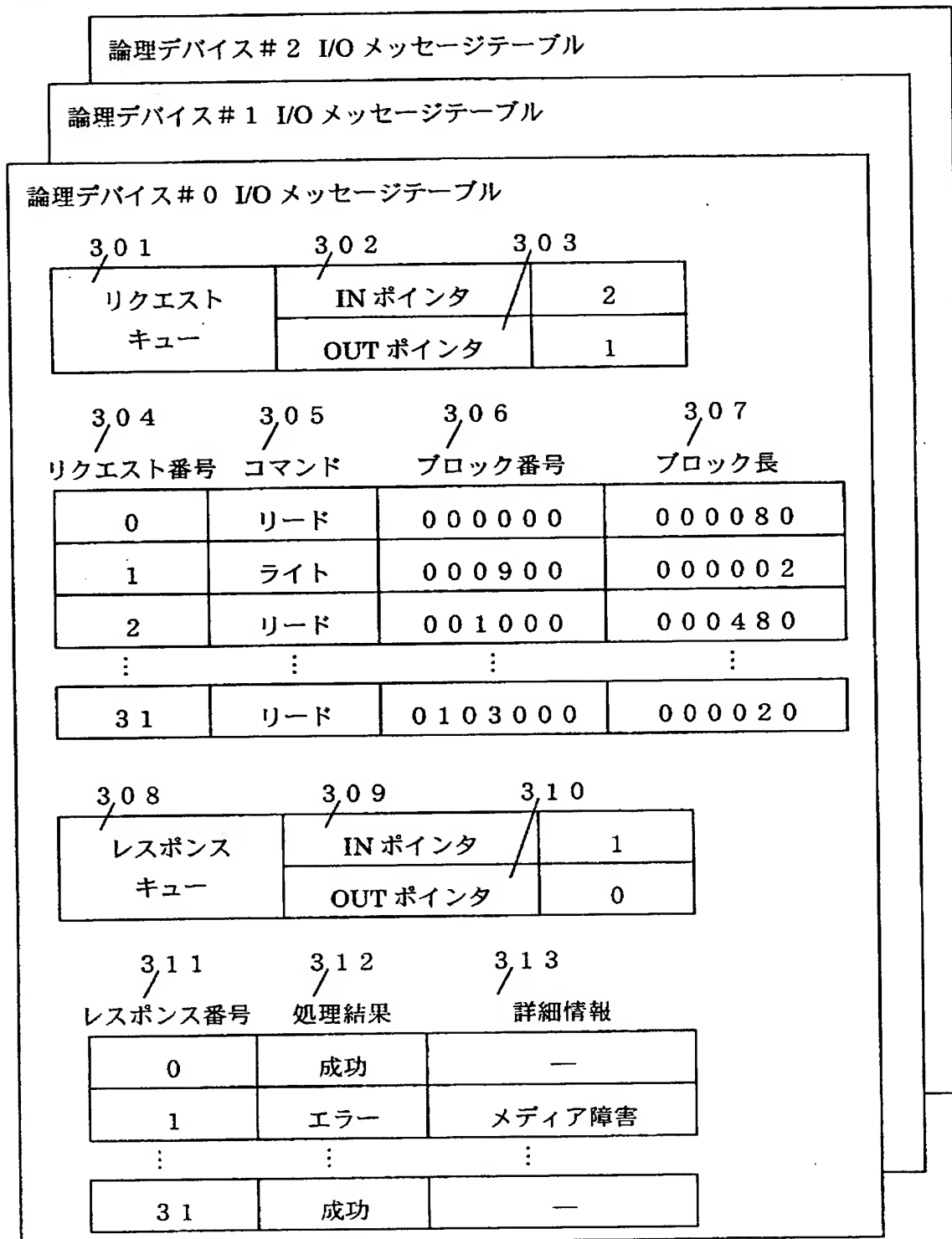
【図 3】

図 3

201 論理 デバイス 番号	202 制御 装置 番号	203 記憶 装置 番号	204 ペア 状態	205 ペア 番号	206 属 性	207 ペア 論理 デバイス 番号	208 コピー ポイン タ
0	0	0	Split	0	正	1	無効
1	1	0	Split	0	副	0	無効
2	1	1	Simplex	無効	正	無効	無効
3	無効	無効	Simplex	無効	正	無効	無効
.
.
.

【図 4】

図 4



【図 5】

図 5

	0	1	2	3	4	5	6	7	...
ペア番号 0	0	0	1	0	1	1	1	0	...
	1	1	1	1	1	1	0	0	...

	0	0	0	0	1	0	0	0	...
ペア番号 1	0	0	0	0	0	0	0	0	...
	0	0	0	0	0	0	0	0	...

	0	0	0	0	0	0	0	0	...
.
.
.
ペア番号 N	0	0	0	0	0	0	0	0	...
	0	0	0	0	0	0	1	1	...

	1	1	1	1	0	0	0	0	...

【図 6】

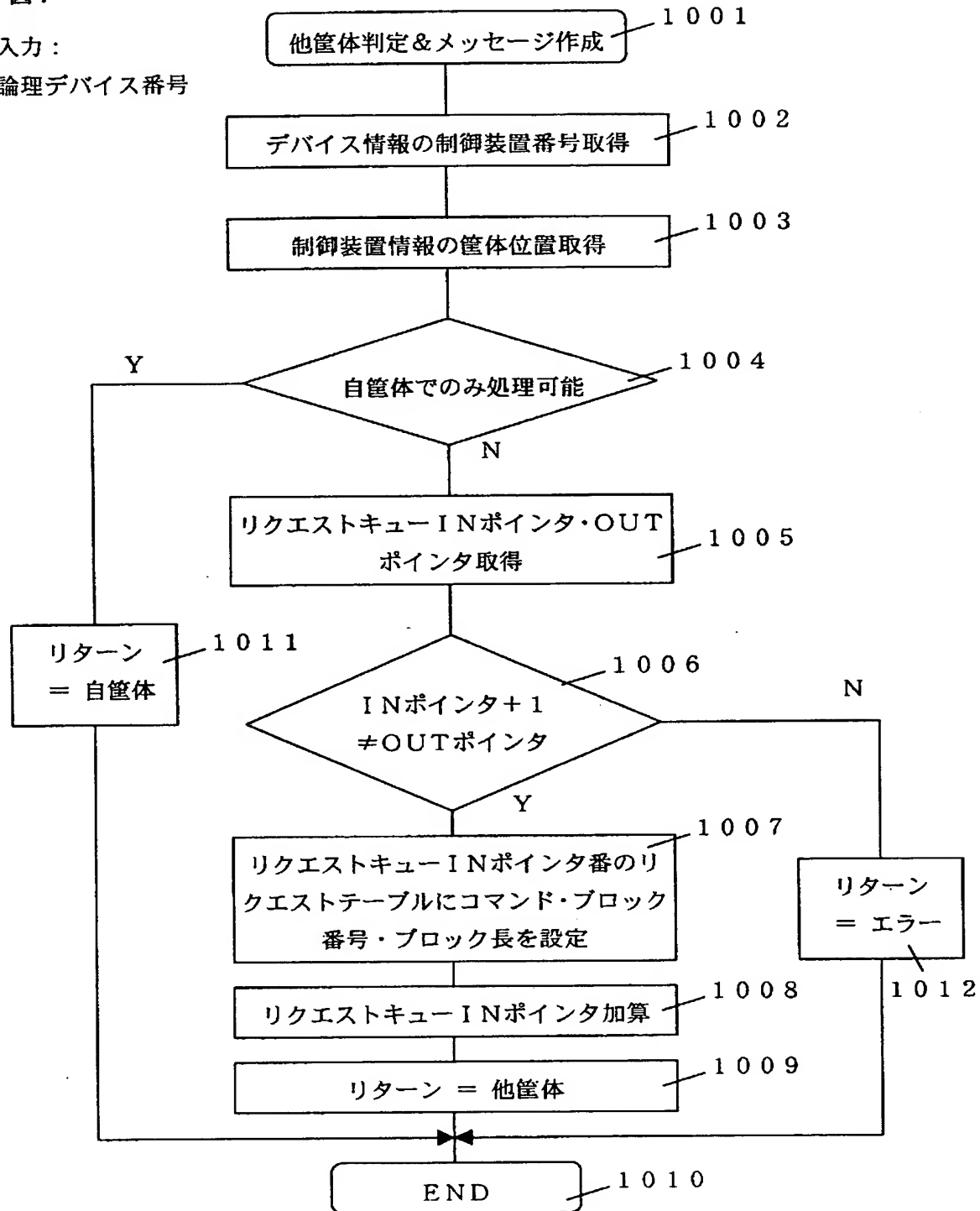
図 6

401 論理 デバイス 番号	402 全 I/O カウンタ	403 リード I/O カウンタ	404 ライト I/O カウンタ
0	0020	0010	0010
1	0010	0006	0004
2	0100	0100	0000
3	0000	0000	0000
・	・	・	・
・	・	・	・
・	・	・	・

【図 7】

図 7

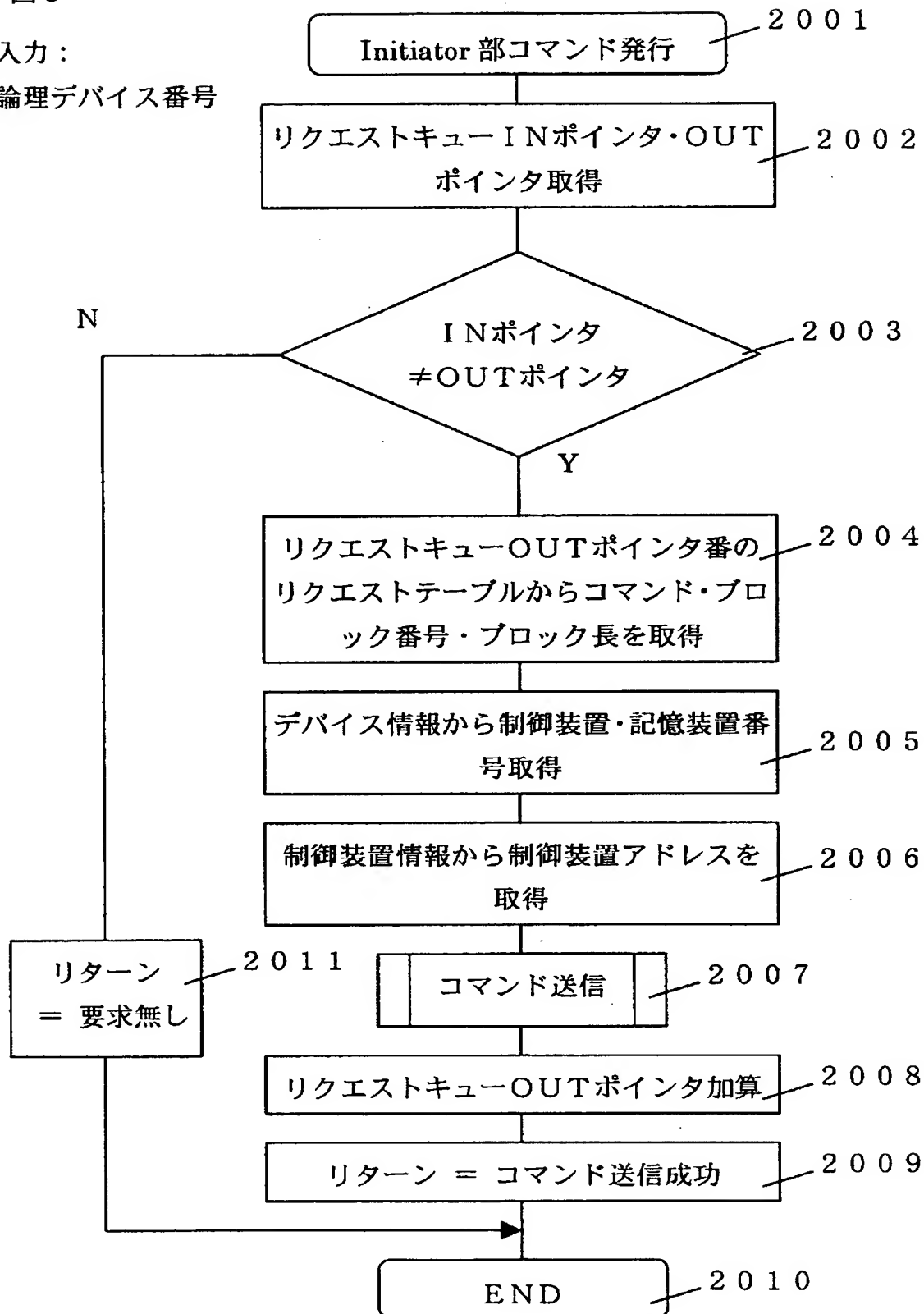
入力：
論理デバイス番号



【図 8】

図 8

入力：
論理デバイス番号



【図 9】

図 9

入力：

論理デバイス番号

更新ブロック番号

更新ブロック長

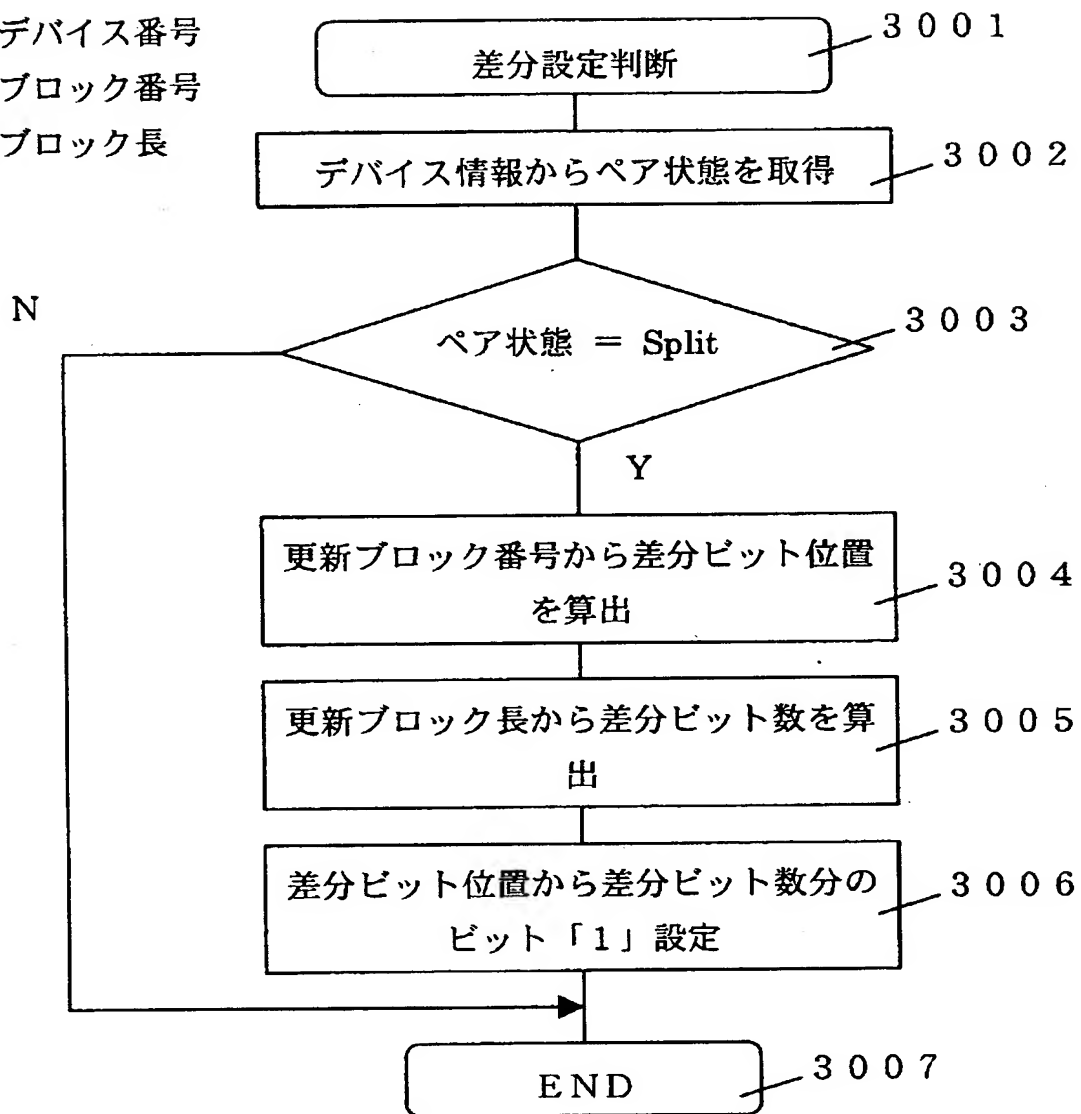


図 10]

図 10

入力：

論理デバイス番号

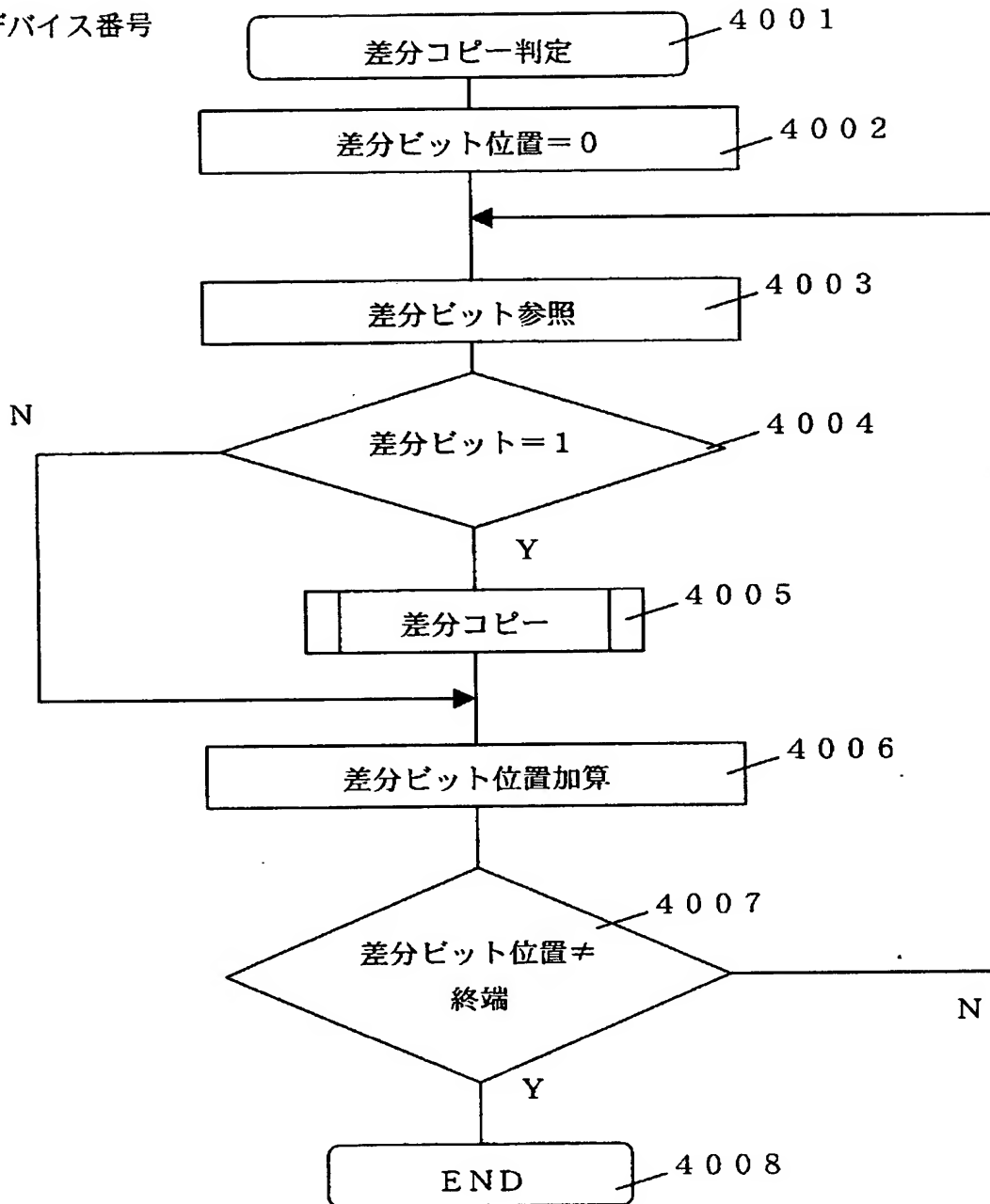
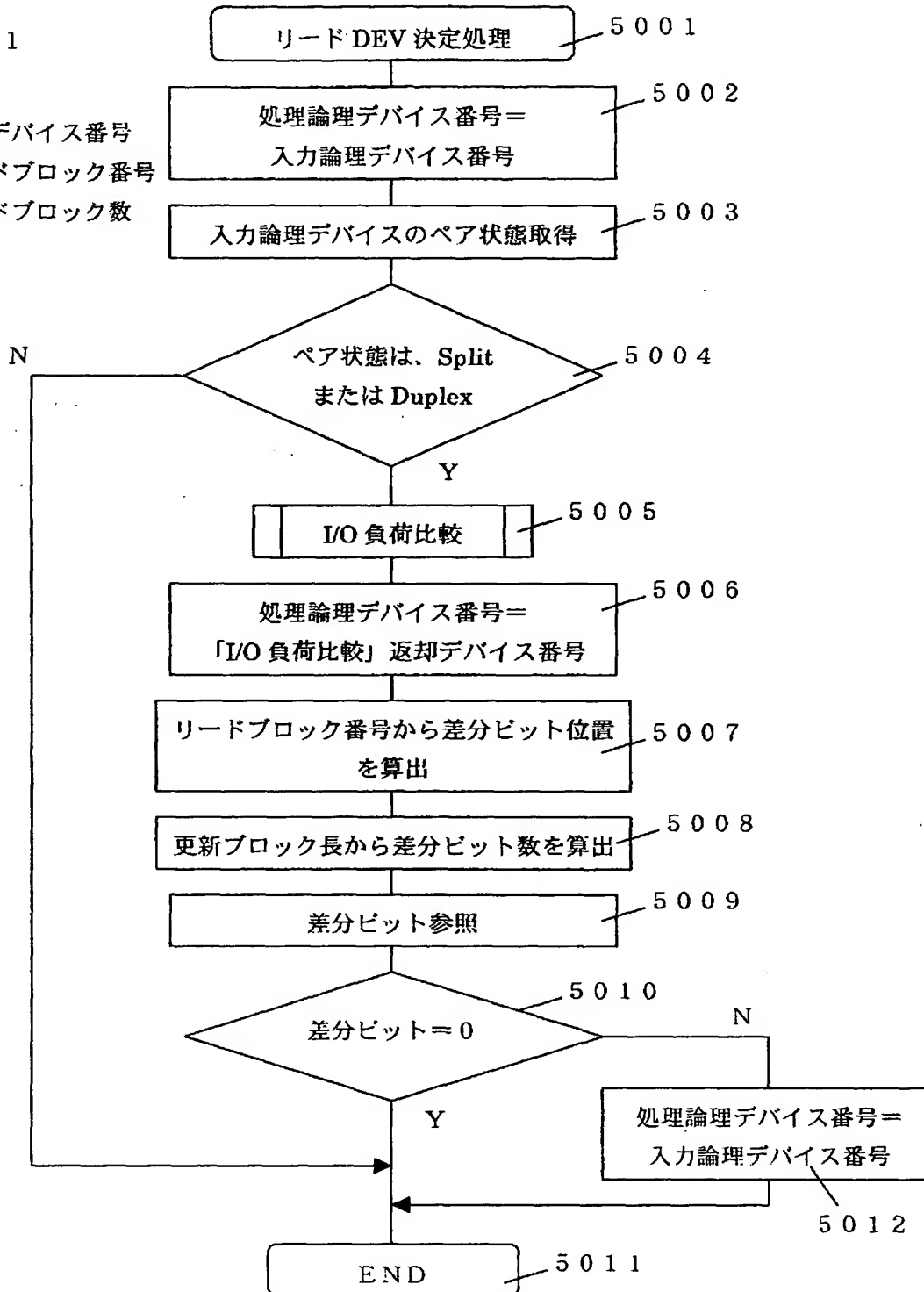


図 11

図 11

入力：
論理デバイス番号
リードブロック番号
リードブロック数

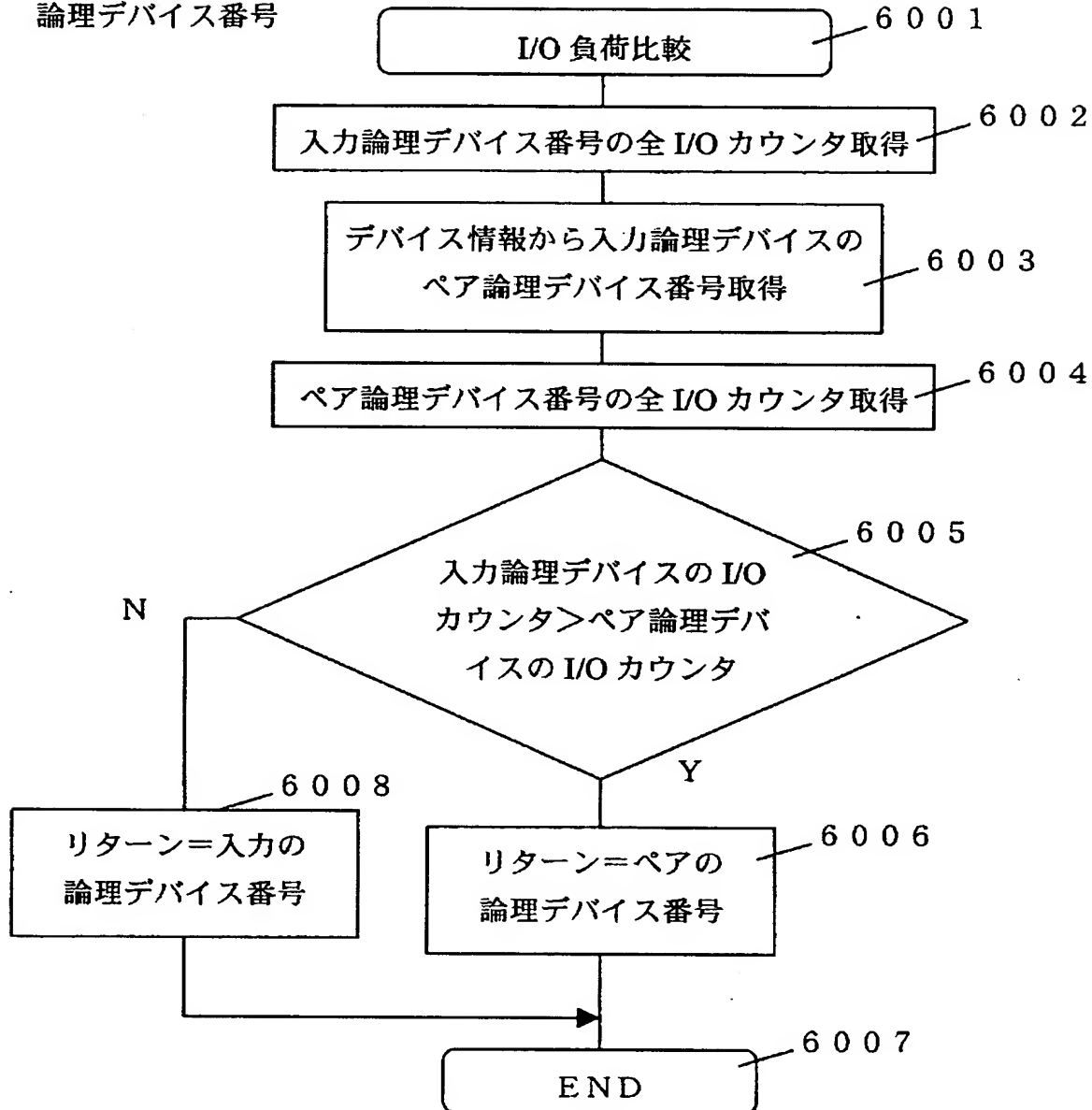


【図 12】

図 12

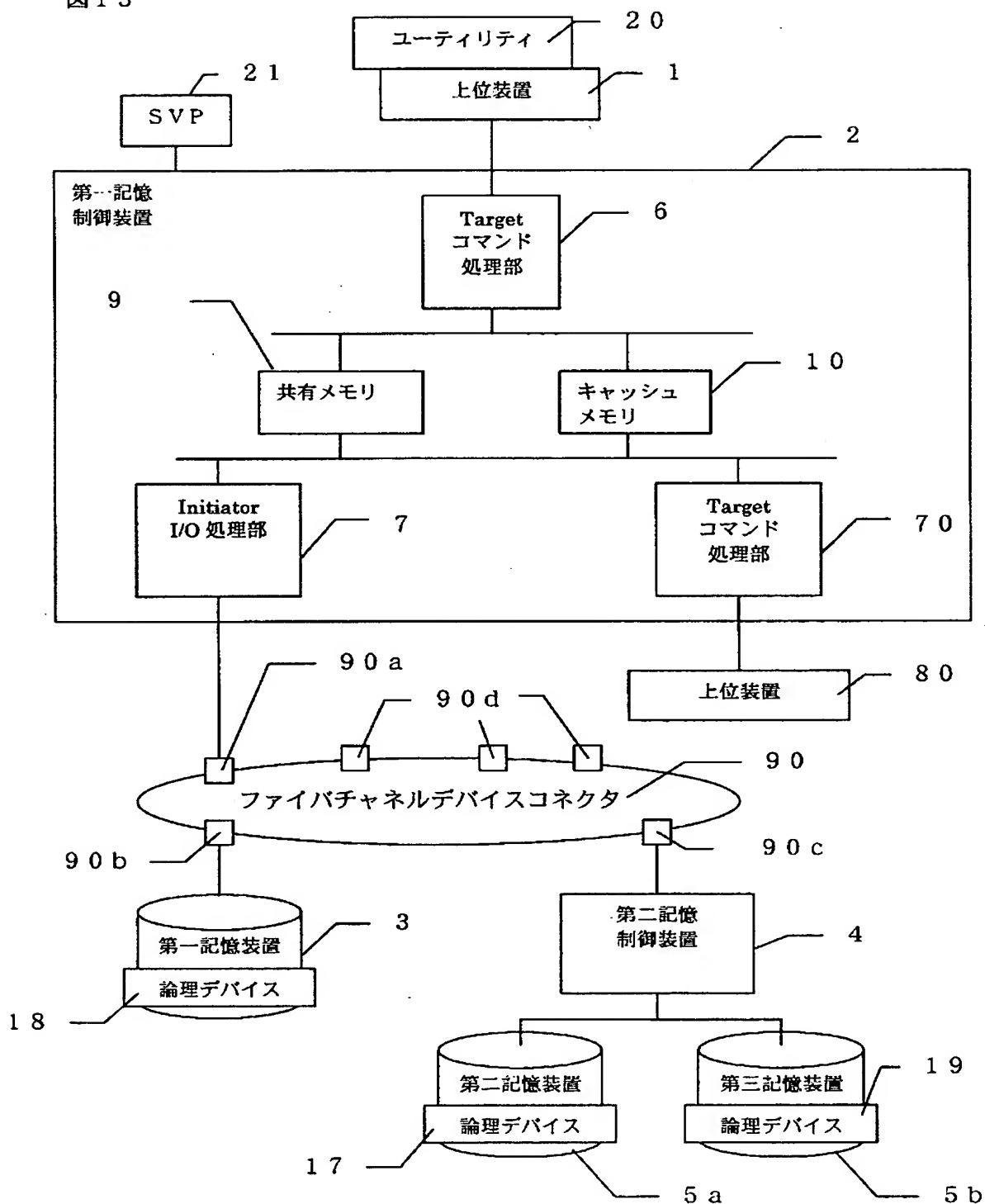
入力：

論理デバイス番号



【図 13】

図 13



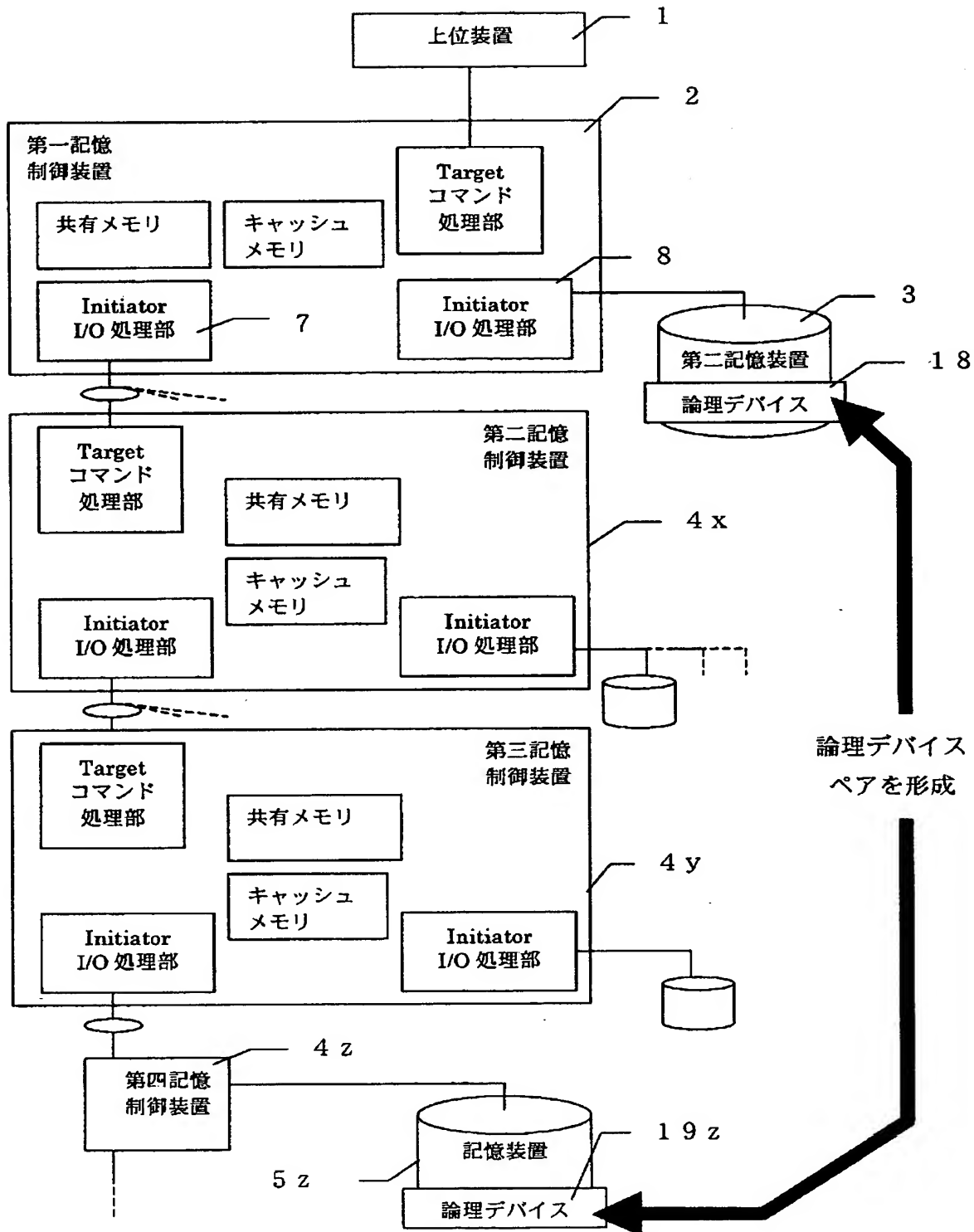
【図 1 4】

図 1 4

1 0 1	1 0 2	1 0 3	1 0 4	1 0 5
制御装置 番号	筐体 位置	イニシエータ 制御部番号	ノード ネーム	アドレス
0	自	0	HTC	0 0 0 0 E 8
1	他	0	SBC	0 0 0 0 E 4
2	無効	無効	無効	無効
.
.
.

【図 15】

図 15



【書類名】 要約書

【要約】

【課題】複数の記憶制御装置を統合的に管理する記憶システムを提供する。

【解決手段】第2の記憶制御装置に対する上位装置からのI/Oを第1の記憶制御装置を介して処理する。第1の記憶制御装置は、第2の記憶制御装置の管理情報を保持することで、第1及び第2の記憶制御装置間においても論理デバイスのペア生成を可能にする。

【効果】第2の記憶制御装置を第1の記憶制御装置から制御することが可能となる。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2003-093118	
受付番号	50300522897	
書類名	特許願	
担当官	第七担当上席	0096
作成日	平成15年 4月 1日	

<認定情報・付加情報>

【提出日】	平成15年 3月31日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所